The Hyper-Morpheme

Language Technology and Text

Ádám Kis
SZAK Publishers Ltd.
adam.kis@szak.hu

Balázs Kis
MorphoLogic Ltd.
kis@morphologic.hu

This paper addresses the representation of the entirety or specific parts and levels of text, from the aspect of their computational treatment. The approach differs from the 'traditional' syntactic or analysis-based aspects of computational linguistics. Instead, it looks at the contents of the text as a complex entity. The authors use three application examples – references (or links), searching and machine translation – to show that the (modelling) problems of computational linguistics all have one common root: namely, the limitations of the computers' pattern matching capabilities. The principal question here is, is one able to overcome these limitations, or is it possible to compare or search for texts based on their information content? This paper raises tasks for researchers without having a solution ready for them.